

REMARKS

Applicants request reconsideration and allowance of the subject application in view of the foregoing amendments and the following remarks.

Claims 1-4, 6, 8, 10, 12, 14, 18, 19, 21, 23, 25, 26, 28, 30, 32, 34, 35, 37, and 39 are pending in this application, with Claims 1, 6, 21, 25, 26, 30, 34, and 35 being independent. Claims 5, 7, 9, 11, 13, 15-17, 20, 22, 24, 27, 29, 31, 33, 36 and 38 have been cancelled without prejudice to or disclaimer of the subject matter contained therein. Claims 1-4, 6, 8, 10, 12, 14, 18, 19, 21, 23, 25, 26, 28, 30, 32, 34, 35, and 37 have been amended. Claim 39 is newly presented. No new matter is believed to have been added.

The title of the invention has been objected to as not being descriptive. Applicants have herein amended the title of the invention and submit that the objection has been overcome. Withdrawal of the objection is requested.

The Summary of the Invention has been objected to as not being a brief summary. Applicants have submitted a replacement summary and request withdrawal of the objection.

The specification has been objected to as failing to provide proper antecedent basis for the claimed subject matter. Specifically, the "data conversion condition" of claims 3, 10, 12, 13, 19, 20, 23, and 24 is said to be unsupported by the specification and so broad a term as to include virtually any situation. These claims have herein been either cancelled or amended such that the objected language is no longer present therein. Applicants request reconsideration and withdrawal of this objection.

The Office Action states that Claim 4 is subject to interpretation. Applicants have amended this claim herein and submit that the language of the claim is clear.

Claim 36 has been objected to as containing an informality. Since this claim has been cancelled herein, Applicants submit that the objection is moot.

Claims 1-9, 14, 15, 18, 21, 22, 25-27, 30, 31, 34, and 35 have been rejected under 35 U.S.C. § 102(e) as being directly anticipated by UK Patent No. GB 2 323 694 A (“Bijl”). Claims 10, 11, 16, 17, 19, 20, 23, 24, 28, 29, 32, 33, 37, and 38 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over Bijl in view of U.S. Patent No. 5,553,119 (“McAllister”). Claims 12 and 13 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over Bijl in view of McAllister, and further in view of U.S. Patent No. 6,263,202 (“Kato”). These rejections are respectfully traversed.

The present invention as recited in the pending independent claims relates to environment adaptation for speech recognition. In a conventional speech recognition system, a speech input terminal transmits inputted speech data to a speech recognition apparatus through a network, and the speech recognition apparatus executes speech recognition for the speech data. In such a system, because users, speech input terminals, and circumstances can all vary, adaptation of the speech recognition to an environment at the side of the speech input terminal is needed. The environment includes, for example, a hardware characteristic of the speech input terminal (such as a microphone characteristic), a noise characteristic, or a speaker characteristic (such as accent).

According to the present invention recited in the independent claims, a speech input terminal creates a model for environment adaptation for speech recognition. The model is

based on information, that is, speech or noise, captured by the terminal. A speech recognition apparatus executes speech recognition based on the model; thus, the speech recognition apparatus is able to adapt an original speech recognition model to an environment at the side of the speech input terminal, using the model from the terminal.

An advantage of this arrangement is that real-time environment adaptation for speech recognition can be achieved. For example, as shown in Fig. 2, a model is created at step 405 prior to speech input at step 412 and speech recognition at step 415, and the created model is reflected in the speech recognition at step 415. Therefore, a speech recognition result that has accounted for an environment at the side of a speech input terminal at that time can be obtained. Additionally, since it is not the captured information itself, but the model that is transmitted from the speech input terminal to a speech recognition apparatus, the amount of data being communicated is reduced.

Independent Claim 1 of the invention, as amended, recites a speech input terminal in a speech communication system including the speech input terminal for transmitting inputted speech data to a speech recognition apparatus through a network, and the speech recognition apparatus executing speech recognition processing for the speech data transmitted from the speech input terminal. The speech input terminal includes speech input means, means for creating a model based on information captured by the speech input means, the model being for environment adaptation for speech recognition, and communication means for transmitting the model to the speech recognition apparatus.

Independent Claim 6, as amended, recites a speech recognition apparatus in a speech communication system (corresponding to a speech communication system as described

above for Claim 1). The speech recognition apparatus includes speech recognition means for execution speech recognition processing for the speech data transmitted from the speech input terminal through the network, and means for receiving a model for environment adaptation for speech recognition from the speech input terminal, the model being created by the speech input terminal based on information captured by the speech input terminal, wherein the speech recognition means executes speech recognition processing on the basis of the model.

Independent Claim 21, as amended, recites a speech communication system (corresponding to a speech communication system as described above for Claim 1). In the system, the speech input terminal includes speech input means, means for creating a model based on information captured by the speech input means, the model being for environment adaptation for speech recognition, and communication means for transmitting the model to the speech recognition apparatus, which receives the model. The speech recognition apparatus includes means for executing speech recognition processing on the basis of the model.

Independent Claims 25, 26, and 30 are method claims reciting features that generally correspond to those recited in Claims 1, 6, and 21, respectively.

Independent Claims 35 and 35 are storage medium claims reciting features that generally correspond to those recited in Claims 1 and 6, respectively.

The primary reference to Bijl discloses environment adaptation for speech recognition in a communication system including a user terminal transmitting inputted speech data to a speech recognition processor through a network. Bijl teaches two general types of environment adaptation.

(i) The speech recognition processor is adapted based on information such as a user's subject matter area, accent, gender, and so on. Data for adaptation is accumulated by pooling the data according, for example, to different accents. This aggregation of data from numerous users is intended to improve performance of the automatic speech recognition processors for subsequent users having an account for which there is pooled data. Identification of the user as belonging to a particular account group allows selection of the acoustic models for that group, which are then applied for the user.

(ii) For future processing, a speech recognition processor is adapted based on correction of a result of the speech recognition. The result can be sent to one of several correction units. In the case of having identified a user as belonging to a certain accent group, the result can be sent to a correction unit in an area where that accent is familiar, or where a particular human corrector is familiar with the user's accent. The correction is effected manually by the human corrector. Adaptation is not reflected in speech recognition for the current user, but for a subsequent user.

Applicants submit that Bijl, however, does not teach or suggest at least the features of the claimed invention, recited among various other features in the independent claims, that a model for environment adaptation for speech recognition, based on information captured at a speech input terminal, is created at the speech input terminal, and that this model (not the information itself) is transmitted to a speech recognition apparatus. The system according to Bijl depends on pools of data, already gathered and aggregated before a particular user begins using the system at any given time, in order to select pre-prepared acoustic models to be applied for the user. In contrast, speech recognition according to the invention is not reliant upon information

previously gathered from other users. Further, whereas the selection of acoustic models according to Bijl occurs at the speech recognition processor side, creation of a model occurs at the side of the speech input terminal, as recited in the independent claims. Information captured at a speech input terminal does not need to be sent from the terminal for processing elsewhere before an appropriate model is returned; rather, a model is created at the terminal itself.

Applicants therefore submit that the independent claims patentably distinguish the invention over Bijl. Accordingly, reconsideration and withdrawal of the § 102 rejection are respectfully requested.

The secondary citation to McAllister relates to recognition of speech signals using caller demographics and/or by actively prompting a user to provide certain responses. An appropriate recognition model or device is then selected. The tertiary citation to Kato relates to a communication system and wireless communication terminal device that allow a user to select the form of output of a message, including emotion, tone color, and language/dialect.

Applicants submit that neither McAllister nor Kato, whether taken alone or in either of the combinations with Bijl proposed in the Office Action, remedies the deficiencies in Bijl discussed above. Specifically, Applicants submit that the references fail to teach or suggest at least that a model for environment adaptation for speech recognition, based on information captured at a speech input terminal, is created at the speech input terminal, and that this model is transmitted to a speech recognition apparatus, as recited in the independent claims. Accordingly, Applicants submit that the independent claims patentably distinguish the invention over McAllister and Kato, taken alone or in the proposed combinations. Reconsideration and withdrawal of the § 103 rejections are respectfully requested.

Applicants submit that the independent claims patentably define the invention over the cited art. Further, the dependent claims should be allowable for the same reasons that the base claims from which they depend are allowable, and further due to the additional features that they recite. Individual consideration of each dependent claim is respectfully requested.

Applicants submit that the application is in condition for allowance. Favorable consideration of the claims and passage to issue of the application at the Examiner's earliest convenience are requested.

Applicants' undersigned attorney may be reached in our Washington, D.C. office by telephone at (202) 530-1010. All correspondence should continue to be directed to our below-listed address.

Respectfully submitted,

  
\_\_\_\_\_  
Attorney for Applicants  
Melody H. Wu  
Registration No. 52,376

FITZPATRICK, CELLA, HARPER & SCINTO  
30 Rockefeller Plaza  
New York, New York 10112-3801  
Facsimile: (212) 218-2200  
MHW:ayr  
151358 v 1